A/No fee

08/319745

PATCHED GENES AND THEIR USE

INTRODUCTION

Technical Field

The field of this invention concerns segment polarity genes and their uses.

Background

Segment polarity genes were discovered in flies as mutations which change the pattern of structures of the body segments. Mutations in the genes cause animals to develop the changed patterns on the surfaces of body segments, the changes affecting the pattern along the head to tail axis. For example, mutations in the gene *patched* cause each body segment to develop without the normal structures in the center of each segment. In their stead is a mirror image of the pattern normally found in the anterior segment. Thus cells in the center of the segment make the wrong structures, and point them in the wrong direction with reference to the over all head-to-tail polarity of the animal. About sixteen genes in the class are known. The encoded proteins include kinases, transcription factors, a cell junction protein, two secreted proteins called wingless (WG) and hedgehog (HH), a single transmembrane protein called *patched* (PTC), and some novel proteins not related to any known protein. All of these proteins are believed to work together in signaling pathways that inform cells about their neighbors in order to set cell fates and polarities.

Many of the segment polarity proteins of *Drosophila* and other invertebrates are closely related to vertebrate proteins, implying that the molecular mechanisms involved are ancient. Among the vertebrate proteins related to the fly genes are En-1 and -2, which act in vertebrate brain development and WNT-1, which is also involved in brain development, but was first found as the oncogene implicated in many cases of mouse breast cancer. In flies, the *patched* gene is transcribed into RNA in a complex and dynamic pattern in embryos, including fine

transverse stripes in each body segment primordium. The encoded protein is predicted to contain many transmembrane domains. It has no significant similarity to any other known protein. Other proteins having large numbers of transmembrane domains include a variety of membrane receptors, channels through membranes and transporters through membranes.

The hedgehog (HH) protein of flies has been shown to have at least three vertebrate relatives: *Sonic hedgehog* (*Shh*); *Indian hedgehog*, and *Desert hedgehog*. The *Shh* is expressed in a group of cells at the posterior of each developing limb bud. This is exactly the same group of cells found to have an important role in signaling polarity to the developing limb. The signal appears to be graded, with cells close to the posterior source of the signal forming posterior digits and other limb structures and cells farther from the signal source forming more anterior structures. It has been known for many years that transplantation of the signaling cells, a region of the limb bud known as the "zone of polarizing activity (ZPA)" has dramatic effects on limb patterning. Implanting a second ZPA anterior to the limb bud causes a limb to develop with posterior features replacing the anterior ones (in essence little fingers instead of thumbs). *Shh* has been found to be the long sought ZPA signal. Cultured cells making *Shh* protein (SHH), when implanted into the anterior limb bud region, have the same effect as an implanted ZPA. This establishes that *Shh* is clearly a critical trigger of posterior limb development.

The factor in the ZPA has been thought for some time to be related to another important developmental signal that polarizes the developing spinal cord. The notochord, a rod of mesoderm that runs along the dorsal side of early vertebrate embryos, is a signal source that polarizes the neural tube along the dorsal-ventral axis. The signal causes the part of the neural tube nearest to the notochord to form floor plate, a morphologically distinct part of the neural tube. The floor plate, in turn, sends

out signals to the more dorsal parts of the neural tube to further determine cell fates. The ZPA was reported to have the same signaling effect as the notochord when transplanted to be adjacent to the neural tube, suggesting the ZPA makes the same signal as the notochord. In keeping with this view, *Shh* was found to be produced by notochord cells and floor plate cells. Tests of extra expression of *Shh* in mice led to the finding of extra expression of floor plate genes in cells which would not normally turn them on. Therefore *Shh* appears to be a component of the signal from notochord to floor plate and from floor plate to more dorsal parts of the neural tube.

Besides limb and neural tubes, vertebrate hedgehog genes are also expressed in many other tissues including, but not limited to the peripheral nervous system, brain, lung, liver, kidney, tooth primordia, genitalia, and hindgut and foregut endoderm.

PTC has been proposed as a receptor for HH protein based on genetic experiments in flies. A model for the relationship is that PTC acts through a largely unknown pathway to inactivate both its own transcription and the transcription of the *wingless* segment polarity gene. This model proposes that HH protein, secreted from adjacent cells, binds to the PTC receptor, inactivates it, and thereby prevents PTC from turning off its own transcription or that of *wingless*. A number of experiments have shown coordinate events between PTC and HH.

Relevant Literature

Descriptions of *patched*, by itself or its role with *hedgehog* may be found in Hooper and Scott, Cell 59, 751-765 (1989); Nakano et al., Nature, 341, 508-513 (1989) (both of which also describes the sequence for *Drosophila patched*) Simcox et al., Development 107, 715-722 (1989); Hidalgo and Ingham, Development, 110, 291-301 (1990); Phillips et al., Development, 110, 105-114 (1990); Sampedro and Guerrero, Nature 353, 187-190 (1991); Ingham

et al., Nature 353, 184-187 (1991); and Taylor et al., Mechanisms of Development 42, 89-96 (1993). Discussions of the role of *hedgehog* include Riddle et al., Cell 75, 1401-1416 (1993); Echelard et al., Cell 75, 1417-1430 (1993); Krauss et al., Cell 75, 1431-1444 (1993); Tabata and Kornberg, Cell 76, 89-102 (1994); Heemskerk & DiNardo, Cell 76, 449-460 (1994); Relink et al., Cell 76, 761-775 (1994); and a short review article by Ingham, Current Biology 4, 347-350 (1994). The sequence for the Drosophila 5′ non-coding region was reported to the GenBank, accession number M28418, referred to in Hooper and Scott (1989), *supra*. See also, Forbes, et al., Development 1993 Supplement 115-124.

SUMMARY OF THE INVENTION

Methods for isolating *patched* genes, particularly mammalian *patched* genes, and the mouse *and inveterbrate patched* genes and sequences are provided. The methods include identification of *patched* genes from other species, as well as members of the same family of proteins. The subject genes provide methods for producing the patched protein, where the genes and proteins may be used as probes for research, diagnosis, binding of *hedgehog* protein for its isolation and purification, gene therapy, as well as other utilities.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a graph having a restriction map of about 10kbp of the 5′ region upstream from the initiation codon of *Drosophila patched* gene and bar graphs of constructs of truncated portions of the 5′ region joined to β-galactosidase, where the constructs are introduced into fly cell lines for the production of embryos. The expression of β-gal in the embryos is indicated in the right-hand table during early and late development of the embryo. The greater the number of +'s, the more intense the staining.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Methods are provided for identifying members of the *patched* (*ptc*) gene family from invertebrate and vertebrate, e.g. mammalian, species, as well as the entire cDNA sequence of the mouse *patched* gene. Also, sequences for invertebrate *patched* genes are provided. The *patched* gene encodes a transmembrane protein having a large number of transmembrane sequences.

In identifying the mouse *patched* gene, primers were employed to move through the evolutionary tree from the known *Drosophila ptc* sequence. Two primers are employed from the *Drosophila* sequence with appropriate restriction enzyme linkers to amplify portions of genomic DNA of a related invertebrate, such as mosquito. The sequences are selected from regions which are not likely to diverge over evolutionary time and are of low degeneracy. Conveniently, the regions are the N-terminal proximal sequence, generally within the first 1.5kb, usually within the first 1kb, of the coding portion of the cDNA, conveniently in the first hydrophilic loop of the protein. Employing the polymerase chain reaction with the primers, a band can be obtained from mosquito genomic DNA. The band may then be amplified and used in turn as a probe. One may use this probe to probe a cDNA library from an organism in a different branch of the evolutionary tree, such as a butterfly. By screening the library and identifying sequences which hybridize to the probe, a portion of the butterfly *patched* gene may be obtained. One or more of the resulting clones may then be used to rescreen the library to obtain an extended sequence, up to and including the entire coding region, as well as the non-coding 5′- and 3′-sequences. As appropriate, one may sequence all or a portion of the resulting cDNA coding sequence.

One may then screen a genomic or cDNA library of a species higher in the evolutionary scale with appropriate probes from one or both of the prior sequences. Of particular interest is screening a genomic library, of a

distantly related invertebrate, e.g. beetle, where one may use a combination of the sequences obtained from the previous two species, in this case, the *Drosophila* and the butterfly. By appropriate techniques, one may identify specific clones which bind to the probes, which may then be screened for cross hybridization with each of the probes individually. The resulting fragments may then be amplified, e.g. by subcloning.

By having all or parts of the 4 different *patched* genes, in the presently illustrated example, *Drosophila* (fly), mosquito, butterfly and beetle, one can now compare the *patched* genes for conserved sequences. Cells from an appropriate mammalian limb bud or other cells expressing *patched*, such as notochord, neural tube, gut, lung buds, or other tissue, particularly fetal tissue, may be employed for screening. Alternatively, adult tissue which produces *patched* may be employed for screening. Based on the consensus sequence available from the 4 other species, one can develop probes where at each site at least 2 of the sequences have the same nucleotide and where the site varies that each species has a unique nucleotide, inosine may be used, which binds to all 4 nucleotides.

Either PCR may be employed using primers or, if desired, a genomic library from an appropriate source may be probed. With PCR, one may use a cDNA library or use reverse transcriptase-PCR, where mRNA is available from the tissue. Usually, where fetal tissue is employed, one will employ tissue from the first or second trimester, preferably the latter half of the first trimester or the second trimester, depending upon the particular host. The age and source of tissue will depend to a significant degree on the ability to surgically isolate the tissue based on its size, the level of expression of *patched* in the cells of the tissue, the accessibility of the tissue, the number of cells expressing *patched* and the like. The amount of tissue available should be large enough so as to provide for a sufficient amount of mRNA to be usefully transcribed and amplified. With mouse tissue, limb bud of

from about 10 to 15 dpc (days post conception) may be employed.

In the primers, the complementary binding sequence will usually be at least 14 nucleotides, preferably at least about 17nucleotides and usually not more than about 30 nucleotides. The primers may also include a restriction enzyme sequence for isolation and cloning. With RT-PCR, the mRNA may be enriched in accordance with known ways, reverse transcribed, followed by amplification with the appropriate primers. (Procedures employed for molecular cloning may be found in Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1988). Particularly, the primers may conveniently come from the N-terminal proximal sequence or other conserved region, such as those sequences where at least five amino acids are conserved out of eight amino acids in three of the four sequences. This is illustrated by the sequences (SEQ ID NO:12) IITPLDCFWEG, (SEQ ID NO:13) LIVGG, and (SEQ ID NO:14) PF.FWEQY. Resulting PCR products of expected size are subcloned and may be sequenced if desired.

The cloned PCR fragment may then be used as a probe to screen a cDNA library of mammalian tissue cells expressing *patched*, where hybridizing clones may be isolated under appropriate conditions of stringency. Again, the cDNA library should come from tissue which expresses *patched*, which tissue will come within the limitations previously decribed. Clones which hybridize may be subcloned and rescreened. The hybridizing subclones may then be isolated and sequenced or may be further analyzed by employing RNA blots and *in situ* hybridizations in whole and sectioned embryos. Conveniently, a fragment of from about 0.5 to 1kbp of the N-terminal coding region may be employed for the Northern blot.

The mammalian gene may be sequenced and as described above, conserved regions identified and used as primers for investigating other species. The N-terminal proximal

region, the C-terminal region or an intermediate region may be employed for the sequences, where the sequences will be selected having minimum degeneracy and the desired level of conservation over the probe sequence.

The DNA sequence encoding PTC may be cDNA or genomic DNA or fragment thereof, particularly complete exons from the genomic DNA, may be isolated as the sequence substntially free of wild-type sequence from the chromosome, may be a 50 kbp fragment or smaller fragment, may be joined to heterologous or foreign DNA, which may be a single nucleotide, oligonucleotide of up to 50bp, which may be a restriction site or other identifying DNA for use as a primer, probe or the like, or a nucleic acid of greater than 50bp, where the nucleic acid may be a portion of a cloning or expression vector, comprise the regulatory regions of an expression cassette, or the like. The DNA may be isolated, purified being substantially free of proteins and other nucleic acids, be in solution, or the like.

The subject gene may be employed for producing all or portions of the *patched* protein. The subject gene or fragment thereof, generally a fragment of at least 12bp, usually at least 18bp, may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into the host. Fragments will usually be immediately joined at the 5′ and/or 3′ terminus to a nucleotide or sequence not found in the natural or wild-type gene, or joined to a label other than a nucleic acid sequence. For expression, an expression cassette may be employed, providing for a transcriptional and translational initiation region, which may be inducible or constitutive, the coding region under the transcriptional control of the transcriptional initiation region, and a transcriptional and translational termination region. Various transcriptional initiation regions may be employed which are functional in the expression host. The peptide may be expressed in prokaryotes or eukaryotes in accordance with conventional ways, depending upon the

purpose for expression. For large production of the protein, a unicellular organism or cells of a higher organism, e.g. eukaryotes such as verterbrates, particularly mammals, may be used as the expression host, such as *E. coli, B, subtilis, S. cerevisiae*, and the like. In many situations, it may be desirable to express the *patched* gene in a mammalian host, whereby the *patched* gene will be transported to the cellular membrane for various studies. The protein has two parts which provide for a total of six transmembrane regions, with a total of six extracellular loops, three for each part. The character of the protein has similarity to a transporter protein. The protein has two conserved glycosylation signal triads.

The subject nucleic acid sequences may be modified for a number of purposes, particularly where they will be used intracellularly, for example, by being joined to a nucleic acid cleaving agent, e.g. a chelated metal ion, such as iron or chromium for cleavage of the gene; as an antisense sequence; or the like. Modifications may include replacing oxygen of the phosphate esters with sulfur or nitrogen, replacing the phosphate with amide, etc.

With the availability of the protein in large amounts by employing an expression host, the protein may be isolated and purified in accordance with conventional ways. A lysate may be prepared of the expression host and the lysate purified used HPLC, exclusion chromatography, gel electrophoresis, affinity chromatography, or other purification technique. The purified protein will generally be at least about 80% pure, preferably at least about 90% pure, and may be up to 100% pure. By pure is intended free of other proteins, as well as cellular debris.

The polypeptide may be used for the production of antibodies, where short fragments provide for antibodies specific for the particular polypeptide, whereas larger fragments or the entire gene allow for the production of

antibodies over the surface of the polypeptide or protein, where the protein may be in its natural conformation.

Antibodies may be prepared in accordance with conventional ways, where the expressed polypeptide or protein may be used as an immunogen, by itself or conjugated to known immunogenic carriers, e.g. KLH, pre-S HBsAg, other viral or eukaryotic proteins, or the like. Various adjuvants may be employed, with a series of injections, as appropriate. For monoclonal antibodies, after one or more booster injections, the spleen may be isolated, the splenocytes immortalized, and then screened for high affinity antibody binding. The immortalized cells, e.g. hybridomas, producing the desired antibodies may then be expanded. For further description, see Monoclonal Antibodies: A Laboratory Manual, Harlow and Lane eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, 1988. If desired, the mRNA encoding the heavy and light chains may be isolated and mutigenized by cloning in *E. coli*, and the heavy and light chains may be mixed to further enhance the affinity of the antibody. The antibodies may find use in diagnostic assays for detection of the presence of the PTC protein on the surface of cells or to inhibit the transduction of signal by the PTC protein ligand by competing for the binding site.

The mouse *patched* gene encodes a protein (SEQ ID NO:10) which has about 38% identical amino acids to fly PTC (SEQ ID NO:6) over about 1,200 amino acids. This amount of conservation is dispersed through much of the protein excepting the C-terminal region. The mouse protein also has a 50 amino acid insert relative to the fly protein.

The butterfly PTC homolog (SEQ ID NO:4) is 1,370 amino acids long and overall has a 50% amino acid identity (72% similarity) to fly PTC (SEQ ID NO:6). With the exception of a divergent C-terminus, this homology is evenly spread across the coding sequence. A 267bp exon from the beetle *patched* gene encodes an 89 amino acid protein fragment which was found to be 44% and 51%

identical to the corresponding regions of fly and butterfly PTC respectively.

The mouse *ptc* message (SEQ ID NO:9) is about 8 kb long and the message is present in low levels as early as 7 dpc, the abundancy increasing by 11 and 15 dpc. Northern blot indicates a clear decrease in the amount of message at 17 dpc. In the adult, PTC RNA is present in high amounts in the brain and lung, as well as in moderate amounts in the kidney and liver. Weak signals are detected in heart, spleen, skeletal muscle and testes.

In mouse embryos, *ptc* mRNA (SEQ ID NO:9) is present at 7 dpc, using *in situ* hybridization. *ptc* is present at high levels along the neural axis of 8.5 dpc embryos. By 11.5 dpc, *ptc* can be detected in developing lung buds and gut, consistent with its Northern profile. In addition, the gene is present at high levels in the ventricular zone of the central nervous system as well as in the zona limitans of the prosencephalon. *ptc* is also strongly transcribed in the perichondrium condensing cartilage of 11.5 and 13.5 dpc limb buds, as well as in the ventral portion of the somites, a region which is prospective sclerotome and eventually forms bone in the vertebral column. PTC is present in a wide range of tissues from endodermal, mesodermal, as well as ectodermal origin, evidencing the fundamental role in many aspects of embryonic development, including the condensation of cartilage, the patterning of limbs, the differentiation of lung tissue, and the generation of neurons.

The *patched* nucleic acid may be used for isolating the gene from various mammalian sources of interest, particularly primate, more particularly human, or from domestic animals, both pet and farm, e.g. lagomorpha, rodentiae, procine, bovine, feline, canine, ovine, equine, etc. By using probes, particularly labeled probes of DNA sequences, of the *patched* gene, one may be able to isolate mRNA or genomic DNA, which may be then used for identifying mutations, particularly associated with genetic diseases, such as spina bifida, limb defects, lung

defects, problems with tooth development, liver and kidney development, peripheral nervous system development, and other sites where a *patched* gene is involved in regulation. The subject probes can also be used for identifying the level of expression in cells associated with the testis to determine the relationship with the level of expression and sperm production.

The gene or fragments thereof may be used as probes for identifying the 5′ non-coding region comprising the transcriptional initiation region, particularly the enhancer regulating the transcription of *patched*. By probing a genomic library, particularly with a probe comprising the 5′ coding region, one can obtain fragments comprising the 5′ non-coding region. If necessary, one may walk the fragment to obtain further 5′ sequence to ensure that one has at least a functional portion of the enhancer. It is found that the enhancer is proximal to the 5′coding region, a portion being in the transcribed sequence and downstream from the promoter sequences. The transcriptional initiation region may be used for many purposes, studying embryonic development, providing for regulated expression of *patched* protein or other protein of interest during embryonic development or thereafter, and in gene therapy.

The gene may also be used for gene therapy, by transfection of the normal gene into embryonic stem cells or into mature cells. A wide variety of viral vectors can be employed for transfection and stable integration of the gene into the genome of the cells. Alternatively, micro-injection may be employed, fusion, or the like for introduction of genes into a suitable host cell. See, for example, Dhawan et al., Science 254, 1509-1512 (1991) and Smith et al., Molecular and Cellular Biology 1990, 3268-3271.

By providing for the production of large amounts of PTC protein, one can use the protein for identifying ligands which bind to the PTC protein. Particularly, one may produce the protein in cells and employ the polysomes

in columns for isolating ligands for the PTC protein. One may incorporate the PTC protein into lipososmes by combining the protein with appropriate lipid surfactants, e.g. phospholipids, cholesterol, etc., and sonicate the mixture of the PTC protein and the surfactants in an aqueous medium. With one or more established ligands, e.g. *hedgehog*, one may use the PTC protein to screen for antagonists which inhibit the binding of the ligand. In this way, drugs may be identified which can prevent the transduction of signals by the PTC protein in normal or abnormal cells.

The PTC protein, particularly binding fragments thereof, the gene encoding the protein, or fragments thereof, particularly fragments of at least about 18 nucleotides, frequently of at least about 30 nucleotides and up to the entire gene, more particularly sequences associated with the hydrophilic loops, may be employed in a wide variety of assays. In these situations, the particular molecules will normally be joined to another molecule, serving as a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, e.g. magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule which provides for detection, in accordance with known procedures. The assays may be used for detecting the presence of molecules which bind to the *patched* gene or PTC protein, in isolating molecules which bind to the *patched* gene, for measuring the amount of *patched*, either as the protein or the message, for identifying molecules which may serve as agonists or antagonists, or the like.

Various formats may be used in the assays. For example, mammalian or invertebrate cells may be designed where the cells respond when an agonist binds to PTC in

the membrane of the cell. An expression cassette may be introduced into the cell, where the transcriptional initiation region of *patched* is joined to a marker gene, such as β-galactosidase, for which a substrate forming a blue dye is available. A 1.5kb fragment that responds to PTC signaling has been identified and shown to regulate expression of a heterologous gene during embryonic development. When an agonist binds to the PTC protein, the cell will turn blue. By employing a competition between an agonist and a compound of interest, absence of blue color formation will indicate the presence of an antagonist. These assays are well known in the literature. Instead of cells, one may use the protein in a membrane environment and determine binding affinities of compounds. The PTC may be bound to a surface and a labeled ligand for PTC enployed. A number of labels have been indicated previously. The candidate compound is added with the labeled ligand in an appropriate buffered medium to the surface bound PTC. After an incubation to ensure that binding has occurred, the surface may be washed free of any non-specifically bound components of the assay medium, particularly any non-specifically bound labeled ligand, and any label bound to the surface determined. Where the label is an enzyme, substrate producing a detectable product may be used. The label may be detected and measured. By using standards, the binding affinity of the candidate compound may be determined.

The availability of the gene and the protein allows for investigation of the development of the fetus and the role *patched* and other molecules play in such development. By employing antisense sequences of the *patched* gene, where the sequences may be introduced in cells in culture, or a vector providing for transcription of the antisense of the *patched* gene introduced into the cells, one can investigate the role the PTC protein plays in the cellular development. By providing for the PTC protein or fragment thereof in a soluble form which can compete with the normal cellular PTC protein for ligand, one can inhibit

the binding of ligands to the cellular PTC protein to see the effect of variation in concentration of ligands for the PTC protein on the cellular development of the host. Antibodies against PTC can also be used to block function, since PTC is exposed on the cell surface.

The subject gene may also be used for preparing transgenic laboratory animals, which may serve to investigate embryonic development and the role the PTC protein plays in such development. By providing for variation in the expression of the PTC protein, employing different transcriptional initiation regions which may be constitutive or inducible, one can determine the developmental effect of the differences in PTC protein levels. Alternatively, one can use the DNA to knock out the PTC protein in embryonic stem cells, so as to produce hosts with only a single functional *patched* gene or where the host lacks a functional *patched* gene. By employing homologous recombination, one can introduce a *patched* gene, which is differentially regulated, for example, is expressed to the development of the fetus, but not in the adult. One may also provide for expression of the *patched* gene in cells or tissues where it is not normally expressed or at abnormal times of development. One may provide for mis-expression or failure of expression in certain tissue to mimic a human disease. Thus, mouse models of spina bifida or abnormal motor neuron differentiation in the developing spinal cord are made available. In addition, by providing expression of PTC protein in cells in which it is otherwise not normally produced, one can induce changes in cell behavior upon binding of ligand to the PTC protein.

Areas of investigation may include the development of cancer treatments. The *wingless* gene, whose transcription is regulated in flies by PTC, is closely related to a mammalian oncogene, *Wnt-1*, a key factor in many cases of mouse breast cancer. Other Wnt family members, which are secreted signalling proteins, are implicated in many aspects of development. In flies, the signaling factor

decapentaplegic, a member of the TGF-beta family of signaling proteins, known to affect growth and development in mammals, is also controlled by PTC. Since members of both the TGF-beta and Wnt families are expressed in mice in places close to overlapping with *patched*, the common regulation provides an opportunity in treating cancer. Also, for repair and regeneration, proliferation competent cells making PTC protein can find use to promote regeneration and healing for damaged tissue, which tissue may be regenerated by transfecting cells of damaged tissue with the *ptc* gene and its normal transcription initiation region or a modified transcription initiation region. For example, PTC may be useful to stimulate growth of new teeth by engineering cells of the gums or other tissues where PTC protein was during an earlier developmental stage or is expressed.

Since Northern blot analysis indicates that *ptc* is present at high levels in adult lung tissue, the regulation of *ptc* expression or binding to its natural ligand may serve to inhibit proliferation of cancerous lung cells. The availability of the gene encoding PTC and the expression of the gene allows for the development of agonists and antagonists. In addition, PTC is central to the ability of neurons to differentiate early in development. The availability of the gene allows for the introduction of PTC into host diseased tissue, stimulating the fetal program of division and/or differentiation. This could be done in conjunction with other genes which provide for the ligands which regulate PTC activity or by providing for agonists other than the natural ligand.

The availability of the coding region for various *ptc* genes from various species, allows for the isolation of the 5′ non-coding region comprising the promoter and enhancer associated with the *ptc* genes, so as to provide transcriptional and post-transcriptional regulation of the *ptc* gene or other genes, which allow for regulation of genes in relation to the regulation of the *ptc* gene. Since the *ptc* gene is autoregulated, activation of the *ptc*

gene will result in activation of transcription of a gene under the transcriptional control of the transcriptional initiation region of the *ptc* gene. The transcriptional initiation region may be obtained from any host species and introduced into a heterologous host species, where such initiation region is functional to the desired degree in the foreign host. For example, a fragment of from about 1.5kb upstream from the initiation codon, up to about 10kb, preferably up to about 5kb may be used to provide for transcriptional initiation regulated by the PTC protein, particularly the *Drosophila* 5'-non-coding region (SEQ ID NO:11).

The following examples are offered by illustration not by way of limitation.

EXPERIMENTAL

Methods and Materials

I. PCR on Mosquito (*Anopheles gambiae*) Genomic DNA:

PCR primers were based on amino acid stretches of fly PTC that were not likely to diverge over evolutionary time and were of low degeneracy. Two such primers (P2R1 (SEQ ID NO:15): GGACGAATTCAARGTNCAYCARYTNTGG, P4R1: (SEQ ID NO:16) GGACGAATTCCYTCCCARAARCANTC, the underlined sequences are Eco RI linkers) amplified an appropriately sized band from mosquito genomic DNA using the PCR. The program conditions were as follows:

94 C 4 min.; 72 C Add Taq;
[49 C 30 sec.; 72 C 90 sec.; 94 C 15 sec] 3 times
[94 C 15 sec.; 50 C 30 sec.; 72 C 90 sec] 35 times
72 C 10 min; 4 C hold

This band was subcloned into the EcoRV site of pBluescript II and sequenced using the USB Sequence kit.

II. Screen of a Butterfly cDNA Library with Mosquito PCR Product

Using the mosquito PCR product (SEQ ID NO:7) as a probe, a 3 day embryonic *Precis coenia* λgt10 cDNA library (generously provided by Sean Carroll) was screened. Filters were hybridized at 65

C overnight in a solution containing 5xSSC, 10% dextran sulfate, 5x Denhardt's, 200 μg/ml sonicated salmon sperm DNA, and 0.5% SDS. Filters were washed in 0.1X SSC, 0.1% SDS at room temperature several times to remove nonspecific hybridization. Of the 100,000 plaques initially screened, 2 overlapping clones, L1 and L2, were isolated, which corresponded to the N terminus of butterfly PTC. Using L2 as a probe, the library filters were rescreened and 3 additional clones (L5, L7, L8) were isolated which encompassed the remainder of the *ptc* coding sequence. The full length sequence of butterfly *ptc* (SEQ ID NO:3) was determined by ABI automated sequencing.

III. Screen of a *Tribolium* (beetle) Genomic Library with Mosquito PCR Product and 900 bp Fragment from the Butterfly Clone

A λgem11 genomic library from *Tribolium casteneum* (gift of Rob Dennell) was probed with a mixture of the mosquito PCR (SEQ ID NO:7) product and BstXI/EcoRI fragment of L2. Filters were hybridized at 55 C overnight and washed as above. Of the 75,000 plaques screened, 14 clones were identified and the SacI fragment of T8 (SEQ ID NO:1), which crosshybridized with the mosquito and butterfly probes, was subcloned into pBluescript.

IV. PCR on Mouse cDNA Using Degenerate Primers Derived from Regions Conserved in the Four Insect Homologues

Two degenerate PCR primers (P4REV: (SEQ ID NO:17) GGACGAATTCYTNGANTGYTTYTGGGA; P22: (SEQ ID NO:18) CATACCAGCCAAGCTTGTCIGGCCARTGCAT) were designed based on a comparison of PTC amino acid sequences from fly (*Drosophila melanogaster*) (SEQ ID NO:6), mosquito (*Anopheles gambiae*)(SEQ ID NO:8), butterfly (*Precis coenia*)(SEQ ID NO:4), and beetle (*Tribolium casteneum*)(SEQ ID NO:2). I represents inosine, which can form base pairs with all four nucleotides. P22 was used to reverse transcribe RNA from 12.5 dpc mouse limb bud (gift from David Kingsley) for 90 min at 37 C. PCR using P4REV(SEQ ID NO:17) and P22(SEQ ID NO:18) was then performed on 1 μl of the resultant cDNA under the following conditions:

94 C 4 min.; 72 C Add Taq;

[94 C 15 sec.; 50 C 30 sec.; 72 C 90 sec.] 35 times

72 C 10 min.; 4 C hold

PCR products of the expected size were subcloned into the TA vector (Invitrogen) and sequenced with the Sequenase Version 2.0 DNA Sequencing Kit (U.S.B.).

Using the cloned mouse PCR fragment as a probe, 300,000 plaques of a mouse 8.5 dpc λgt10 cDNA library (a gift from Brigid Hogan) was screened at 65°C as above and washed in 2x SSC, 0.1% SDS at room temperature. 7 clones were isolated, and three (M2 M4, and M8) were subcloned into pBluescript II. 200,000 plaques of this library were rescreened using first, a 1.1 kb EcoRI fragment from M2 to identify 6 clones (M9-M16 and secondly a mixed probe containing the most N terminal (XhoI fragment from M2) and most C terminal sequences (BamHI/BglII fragment from M9) to isolate 5 clones (M17-M21). M9, M10, M14, and M17-21 (SEQ ID NO:9) were subcloned into the EcoRI site of pBluescript II (Strategene).

V. RNA Blots and in situ Hybridizations in Whole and Sectioned Mouse Embryos

Northerns:

A mouse embryonic Northern blot and an adult multiple tissue Northern blot (obtained from Clontech) were probed with a 900 bp EcoRI fragment from an N terminal coding region of mouse *ptc*. Hybridization was performed at 65 C in 5x SSPE, 10x Denhardt's, 100 μg/ml sonicated salmon sperm DNA, and 2% SDS. After several short room temperature washes in 2x SSC, 0.05% SDS, the blots were washed at high stringency in 0.1X SSC, 0.1% SDS at 50C.

In situ hybridization of sections:

7.75, 8.5, 11.5, and 13.5 dpc mouse embryos were dissected in PBS and frozen in Tissue-Tek medium at -80 C. 12-16 μm frozen sections were cut, collected onto VectaBond (Vector Laboratories) coated slides, and dried for 30-60 minutes at room temperature. After a 10 minute fixation in 4% paraformaldehyde in PBS, the slides were washed 3 times for 3 minutes in PBS, acetylated for 10 minutes in 0.25% acetic anhydride in triethanolamine, and washed three more times for 5 minutes in PBS. Prehybridization (50% formamide, 5X SSC, 250

μg/ml yeast tRNA, 500 μg/ml sonicated salmon sperm DNA, and 5x Denhardt's) was carried out for 6 hours at room temperature in 50% formamide/5x SSC humidified chambers. The probe, which consisted of 1 kb from the N-terminus of *ptc*, was added at a concentration of 200-1000 ng/ml into the same solution used for prehybridization, and then denatured for five minutes at 80 C. Approximately 75 μl of probe were added to each slide and covered with Parafilm. The slides were incubated overnight at 65 C in the same humidified chamber used previously. The following day, the probe was washed successively in 5X SSC (5 minutes, 65 C), 0.2X SSC (1 hour, 65 C), and 0.2X SSC (10 minutes, room temperature). After five minutes in buffer B1 (0.1M maleic acid, 0.15 M NaCl, pH 7.5), the slides were blocked for 1 hour at room temperature in 1% blocking reagent (Boerhinger-Mannheim) in buffer B1, and then incubated for 4 hours in buffer B1 containing the DIG-AP conjugated antibody (Boerhinger-Mannheim) at a 1:5000 dilution. Excess antibody was removed during two 15 minute washes in buffer B1, followed by five minutes in buffer B3 (100 mM Tris, 100mM NaC1, 5mM $MgCl_2$, pH 9.5). The antibody was detected by adding an alkaline phosphatase substrate (350 μl 75 mg/ml X-phosphate in DMF, 450 μl 50 mg/ml NBT in 70% DMF in 100 mls of buffer B3) and allowing the reaction to proceed over-night in the dark. After a brief rinse in 10 mM Tris, 1mM EDTA, pH 8.0, the slides were mounted with Aquamount (Lerner Laboratories).

VI. *Drosophila* 5-transcriptional initiation region β-gal constructs.

A series of constructs were designed that link different regions of the *ptc* promoter from *Drosophila* to a *LacZ* reporter gene in order to study the cis regulation of the *ptc* expression pattern. See Fig. 1. A 10.8kb BamHI/BspM1 fragment comprising the 5'-non-coding region of the mRNA at its 3'-terminus was obtained and truncated by restriction enzyme digestion as shown in Fig. 1. These expression cassettes were introduced into *Drosophila* lines using a P-element vector (Thummel et al., Gene 74, 445-456 (1988), which were injected into embryos, providing flies which could be grown to produce embryos. (See Spradling and Rubin, Science (1982) 218, 341-347 for a description of the procedure.) The vector used a pUC8 background into which was introduced the white gene to provide for yellow eyes, portions of the P-element for integrtion, and the

constructs were inserted into a polylinker upstream from the *LacZ* gene. The resulting embryos were stained using antibodies to LacZ protein conjugated to HRP and the embryos developed with OPD dye to identify the expression of the *LacZ* gene. The staining pattern is described in Fig. 1, indicating whether there was staining during the early and late development of the embryo.

RESULTS

Isolation of a Mouse *ptc* Gene

Homologues of fly *ptc* (SEQ ID NO:6) were isolated from three insects: mosquito, butterfly and beetle, using either PCR or low stringency library screens. PCR primers to six amino acid stretches of PTC of low mutatability and degeneracy were designed. One primer pair, P2 and P4, amplified an homologous fragment of PTC from mosquito genomic DNA that corresponded to the first hydrophilic loop of the protein. The 348bp PCR product (SEQ ID NO:7) was subcloned and sequenced and when aligned to fly PTC, showed 67% amino acid identity.

The cloned mosquito fragment was used to screen a butterfly lGT 10 cDNA library. Of 100,000 plaques screened, five overlapping clones were isolated and used to obtain the full length coding sequence. The butterfly PTC homologue (SEQ ID NO:4) is 1,370 amino acids long and overall has 50% amino acid identity (72% similarity) to fly PTC. With the exception of a divergent C-terminus, this homology is evenly spread across the coding sequence. The mosquito PCR clone (SEQ ID NO:7) and a corresponding fragment of butterfly cDNA were used to screen a beetle lgem11 genomic library. Of the plaques screened, 14 clones were identified. A fragment of one clone (T8), which hybridized with the original probes, was subcloned and sequenced. This 3kb piece contains an 89 amino acid exon (SEQ ID NO:2) which is 44% and 51% identical to the corresponding regions of fly and butterfly PTC respectively.

Using an alignment of the four insect homologues in the first hydrophilic loop of the PTC, two PCR primers were designed to a five and six amino acid stretch which were identical and of low degeneracy. These primers were used to isolate the mouse homologue using RT-PCR on embryonic limb bud RNA. An appropriately sized band was amplified and upon cloning and sequencing, it was

found to encode a protein 65% identical to fly PTC. Using the cloned PCR product and subsequently, fragments of mouse *ptc* cDNA, a mouse embryonic lcDNA library was screened. From about 300,000 plaques, 17 clones were identified and of these, 7 form overlapping cDNA's which comprise most of the protein-coding sequence (SEQ ID NO:9) .

Comparison of Mouse, Fly and Butterfly Sequences

The deduced mouse PTC protein sequence (SEQ ID NO:10) has about 38% identical amino acids to fly PTC over about 1,200 amino acids. This amount of conservation is dispersed through much of the protein excepting the C-terminal region. The mouse protein also has a 50 amino acid insert relative to the fly protein. Based on the sequence conservation of PTC and the functional conservation of *hedgehog* between fly and mouse, one concludes that *ptc* functions similarly in the two organisms. A comparison of the amino acid sequences of mouse (mptc) (SEQ ID NO:10), butterfly (bptc)(SEQ ID NO:4) and drosophila (ptc) (SEQ ID NO:6) is shown in Table 1.

TABLE 1

```
          mptc
          MASAGNXRRGPGQAGRRREAQTDRGTAPR
          ptc
          MDRDSLPRVPDTHGDVVDE--
          bptc
          MVAPDSEAPSNPRITAAHESPCATEA
                                                                *
          mptc  RAGPGLSAPAQLLRRAFALEQISKGKATGRKAPLWLRAKFQRLLFKLGCYIQK
          NCGKFLV
          ptc   KLFSDLYIRTSWVDAQVALDQIDKGKARGSRTAIYLRSVFQSHLETLGSSVQK
          HAGKVLF
          bptc  RHSADLYIRTSWVDAALALSELEKGNIEGGRTSLWIRAWLQEQLFILGCFLQG
          DAGKVLF
                .    *   .  .    ** .. **.  * .... .*. .*  *  **  .* . **
          *
          mptc  VGLLIFGAFAVGLKAANLETNVEELWVEVGGRVSRELNYTRQKIGEEAMFNPQ
          LMIQTPK
          ptc   VAILVLSTFCVGLKSAQIHSKVHQLWIQEGGRLEAELAYTQKTIGEDESATHQ
          LLIQTTH
          bptc  VAILVLSTFCVGLKSAQIHTRVDQLWVQEGGRLEAELKYTAQALGEADSSTHQ
          LVIQTAK
                *..*....* ****.*.... *..**.. ***.  ** ** . .**      **.***
          mptc  EEGANVLTTEALLQHLDSALQASRVHVYMYNRQWKLEHLCYKSGELITET-GY
          MDQIIEY
          ptc   DPNASVLHPQALLAHLEVLVKATAVKVHLYDTEWGLRDMCNMPSTPSFEGIYY
          IEQILRH
          bptc  DPDVSLLHPGALLEHLKVVHAATRVTVHMYDIEWRLKDLCYSPSIPDFEGYHH
          IESIIDN
                .   ..*   *** **     *. * * .*. .* * ..*  ..    *    .. *.
          mptc  LYPCLIITPLDCFWEGAKLQSGTAYLLGKPPLR----WTNFDPLEFLEELK-- ----
          KIN
```

```
ptc   LIPCSIITPLDCFWEGSQLL-GPESAVVIPGLNQRLLWTTLNPASVMQYMKQK
MSEEKIS
bptc  VIPCAIITPLDCFWEGSKLL-GPDYPIYVPHLKHKLQWTHLNPLEVVEEVK-K L---
KFQ
      . ** ***********..*  *    .  * *     ** ..*   .. .*
*.
mptc  YQVDSWEEMLNKAEVGHGYMDRPCLNPADPDCPATAPNKNSTKPLDVALVLNG
GCQGLSR
ptc   FDFETVEQYMKRAAIGSGYMEKPCLNPLNPNCPDTAPNKNSTQPPDVGAILSG
GCYGYAA
bptc  FPLSTIEAYMKRAGITSAYMKKPCLDPTDPHCPATAPNKKSGHIPDVAAELSH
GCYGFAA
      .   . *  ...* .  .** .***.* .*.** *****.*    **.  *. ** *
.
mptc  KYMHWQEELIVGGTVKNATGKLVSAHALQTMFQLMTPKQMYEHFRGYDYVSHI
NWNEDRA
ptc   KHMHWPEELIVGGRKRNRSGHLRKAQALQSVVQLMTEKEMYDQWQDNYKVHHL
GWTQEKA
bptc  AYMHWPEQLIVGGATRNSTSALRKARXLQTVVQLMGEREMYEYWADHYKVHQI
GWNQEKA
        *** *.*****  .* .. *  *. **.. ***  ..**.   .   * .. *
...*
mptc  AAILEAWQRTYVEVVHQSVAPNSTQK----VLPFTTTTLDDILKSFSDVSVIR
VASGYLL
ptc   AEVLNAWQRNFSREVEQLLRKQSRIATNYDIYVFSSAALDDILAKFSHPSALS
IVIGVAV
bptc  AAVLDAWQRKFAAEVRKI-TTSGSVSSAYSFYPFSTSTLNDILGKFSEVSLKN
IILGYMF
      * .*.**** .   * .     .          *....*.***  **. *   .  *
mptc       MLAYACLTMLRW-DCSKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATT
QVLPFLA
 ptc   TVLYAFCTLLRWRDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNAAST
QVVPFLA
bptc  MLIYVAVTLIQWRDPIRSQAGVGIAGVLLLSITVAAGLGFCALLGIPFNASST
QIVPFLA
      . *   *...* *  ..*..**.*****. .. *****. .*.**
***..**..****
mptc  LGVGVDDVFLLAHAFSETGQNKRIPFEDRTGECLKRTGASVALTSISNVTAFF
MAALIPI
ptc   LGLGVDHIFMLTAAYAESN------RREQTKLILKKVGPSILFSACSTAGSF
FAAAFIPV
bptc  LGLGVQDMFLLTHTYVEQAGD--VPREERTGLVLKKSGLSVLLASLCNVMAFL
AAALLPI
      **.**...*.*. .. *          ..*   **. * *. ...     .*.
**..*.
mptc  PALRAFSLQAAVVVVFNFAMVLLIFPAILSMDLYRREDRRLDIFCCLTSPCVS
RVIQVEP
ptc   PALKVFCLQAAIVMCSNLAAALLVFPAMISLDLRRRTAGRADIFCCCF-PVWK
EQPKVAP
bptc  PAFRVFCLQAAILLLFNLGSILLVFPAMISLDLRRRSAARADLLCCLM-P--- ESP-
---
      **.. * ****...  *..  **.***..*.** **   * *..**   *
mptc  QAYTEPHSNTRYSPPPPYTSHSFAHETHITMQSTVQLRTEYDPHTHVYYTTAE
PRSEISV
ptc   PVLPLNNNNGR---------------------------------------GARH
PKSCNNN
bptc  --LPKKKIPER-------------------------------------AKT
RKNDKTH
                 *                                          .  ..
mptc  QPVTVTQDNLSCQSPESTSSTRDLLSQFSDSSLHCLEPPCTKWTLSSFAEKHY
APFLLKP
ptc   RVPLPAQNPLLEQRADIPGSS-----------HSLASF----SLATFAFQHY
TPFLMRS
bptc  RID-TTRQPLDPDVSENVTKT-----------CCL-SV----SLTKWAKNQY
APFIMRP
      .   ... *  . ..   .             * .    .*.  *
..*.**....
```

```
mptc  KAKVVVILLFLGLLGVSLYGTTRVRDGLDLTDIVPRETREYDFIAAQFKYFSF
YNMYIVT
ptc   WVKFLTVMGFLAALISSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLFGF
YSMYAVT
bptc  AVKVTSMLALIAVILTSVWGATKVKDGLDLTDIVPENTDEHEFLSRQEKYFGF
YNMYAVT
        *   .. ... .  *. ..*...****. *.** .. *  *.  * . *.**.**
**
mptc  QKA-DYPNIQHLLYDLHKSFSNVKYVMLEENKQLPQMWLHYFRDWLQGLQDAF
DSDWETG
ptc   QGNFEYPTQQQLLRDYHDSFVRVPHVIKNDNGGLPDFWLLLFSEWLGNLQKIF
DEEYRDG
bptc  QGNFEYPTNQKLLYEYHDQFVRIPNIIKNDNGGLTKFWLSLFRDWLLDLQVAF
DKEVASG
      *   .**  * ** . *  *  .  .. ..*  *   **  * .**  **  ** .
*
mptc  RIMPNN-YKNGSDDGVLAYKLLVQTGSRDKPIDISQLTK-QRLVDADGIINPS
AFYIYLT
ptc   RLTKECWFPNASSDAILAYKLIVQTGHVDNPVDKELVLT-NRLVNSDGIINQR
AFYNYLS
bptc  CITQEYWCKNASDEGILAYKLMVQTGHVDNPIDKSLITAGHRLVDKDGIINPK
AFYNYLS
       .  .    *.* ...*****.****  *.*.*   .   .***. *****  ***
**.
mptc  AWVSNDPVAYAASQXNIRPHRPEWVHDKADYMPETRLRIPAAEPIEYAQFPFY
LNGLRDT
ptc   AWATNDVFAYGASQGKLYPEPRQYFHQPNEY----DLKIPKSLPLVYAQMPFY
LHGLTDT
bptc  AWATNDALAYGASQGNLKPQPQRWIHSPEDV----HLEIKKSSPLIYTQLPFY
LSGLSDT
      ** .**  **.*** .. *.     *   .      * *  . *. *.* **** **
**
mptc  SDFVEAIEKVRVICNNYTSLGLSSYPNGYPFLFWEQYISLRHWLLLSISVVLA
CTFLVCA
ptc   SQIKTLIGHIRDLSVKYEGFGLPNYPSGIPFIFWEQYMTLRSSLAMILACVLL
AALVLVS
bptc  XSIKTLIRSVRDLCLKYEAKGLPNFPSGIPFLFWEQYLYLRTSLLLALACALA
AVFIAVM
        .   *  .* .  .* . **...*.* **.*****. **  * . ..  *   ..
mptc  VFLLNPWTAGIIVMVLALMTVELFGMMGLIGIKLSAVPVVILIASVGIGVEFT
VHVALAF
ptc   LLLLSVWAAVLVILSVLASLAQIFGAMTLLGIKLSAIPAVILILSVGMMLCFN
VLISLGF
bptc  VLLLNAWAAVLVTLALATLVLQLLGVMALLGVKLSAMPAVLLVLAIGRGVHFT
VHLCLGF
      ..**. *.* .. . .     ...* * *.*.****.* *.*. ..*  . * * .
*.*
mptc  LTAIGDKNHRAMLALEHMFAPVLDGAVSTLLGVLMLAGSEFDFIVRYFFAVLA
ILTVLGV
ptc   MTSVGNRQRRVQLSMQMSLGPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLL
VVLCVGA
bptc  VTSIGCKRRRASLALESVLAPVVHGALAAALAASMLAASECGFVARLFLRLLL
DIVFLGL
      .*..* . .*  *...  ..*...* ... ..  **. *   *. * *  .*  .
.*
mptc  LNGLVLLPVLLSFFGPCPEVSPANGLNRLPTPSPEPPPSVVR----FAVPPGH
TNNGS--
ptc   CNSLLVFPILLSMVGPEAELVPLEHPDRISTPSPLPVRSSKRSGKSYVVQGSR
SSRGSCQ
bptc  IDGLLFFPIVLSILGPAAEVRPIEHPERLSTPSPKCSPIHPRKSSSSSGGGDK
SSRTS--
       ..*. .*..**  ** .*. * .  .*..****       *           ..  *
mptc  --DSSDSEYSSQTTVSGISEELRQYEAQQGAGGPAHQVIVEATENPVFARS-T
VVHPDSR
ptc   KSHHHHHKDLNDPSLTTITEEPQSWKSSNSSIQMPNDWTYQPREQ--RPASYA
APPPAYH
bptc  KSAPRPC----APSLTTITEEPSSWHSSAHSVQSSMQSIVVQPEVVVETTTYN
GSDSASG
```

```
                    ... *.**     .   .   . .       *       .       .
mptc  HQPPLTPRQQPHLDSGSLSPGRQGQQPRRDPP-REGLRPPPYRPRRDAFEIST
EGHSGPS
ptc   KAAAQQHHQHQGPPTTPPPPFPTA-----------------YPPELQSIVVQP
EVTVETT
bptc  RSTPTKSSHGGAITTTKVTATANIKVEVVTPSDRKSRRSYHYYDRRRDRDEDR
DRDRERD
          .   .   .   .                         *               .
mptc  ---------NRDRSGPVGPVLTTLGTQRPPPWAA----LCPATASPSPL
ptc   HS----------DSNTTKVTATANIKVELAMPGRAVRSYNFTS-----
bptc  RDRDRDRDRDRDRDRDRDRERSRERDRRDRYRDERDHRASPREKRQRFW
                              .   .
```

The identity of ten other clones recovered from the mouse library is not determined. These cDNAs cross-hybridize with mouse *ptc* sequence, while differing as to their restriction maps. These genesencode a family of proteins related to the *patched* protein.

Developmental and Tissue Distribution of Mouse PTC RNA

In both the embryonic and adult Northern blots, the *ptc* probe detects a single 8kb message. Further exposure does not reveal any additional minor bands. Developmentally, *ptc* mRNA is present in low levels as early as 7 dpc and becomes quite abundant by 11 and 15 dpc. While the gene is still present at 17 dpc, the Northern blot indicates a clear decrease in the amount of message at this stage. In the adult, *ptc* RNA is present in high amounts in the brain and lung, as well as in moderate amounts in the kidney and liver. Weak signals are detected in heart, spleen, skeletal muscle, and testes.

In situ Hybridization of Mouse PTC in Whole and Section Embryos

Northern analysis indicates that *ptc* mRNA is present at 7 dpc, while there is no detectable signal in sections from 7.75 dpc embryos. This discrepancy is explained by the low level of transcription. In contrast, *ptc* is present at high levels along the neural axis of 8.5 dpc embryos. By 11.5 dpc, *ptc* can be detected in the developing lung buds and gut, consistent with its adult Northern profile. In addition, the gene is present at high levels in the ventricular zone of the central nervous system, as well as in the zona limitans of the prosencephalon. *ptc* is also strongly transcribed in the condensing cartilage of 11.5 and 13.5 dpc limb buds, as well as in the ventral portion of the somites, a region which is prospective sclerotome and eventually forms bone in the vertebral column. *ptc* is present in a wide range of tissues from endodermal,

mesodermal and ectodermal origin supporting its fundamental role in embryonic development.

In accordance with the subject invention, the mammalian *patched* gene is provided which allows for high level production of the *patched* protein, which can serve many purposes. The *patched* protein may be used in a screening for agonists and antagonists, for isolation of its ligand, particularly *hedgehog*, more particularly *Sonic hedgehog* and for assaying for the transcription of the mRNA *ptc*. The protein or fragments thereof may be used to produce antibodies specific for the protein or specific epitopes of the protein. In addition, the gene may be employed for investigating embryonic development, by screening fetal tissue, preparing transgenic animals to serve as models, and the like.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.